

Expected n -Step Product for Gaussian Tours

STEVEN FINCH

December 17, 2015

ABSTRACT. Supplements to Mehta & Normand (1997) are given, with regard to integrals involving Euclidean distances between $n + 1$ random points in d -dimensional space, each visited once.

Let $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_{n+1}$ be independent random Gaussian points in \mathbb{R}^d , all of which have mean vector zero and covariance matrix identity. What can be said about

$$\mu_{d,n} = \mathbb{E}(|\vec{r}_2 - \vec{r}_1| \cdot |\vec{r}_3 - \vec{r}_2| \cdot |\vec{r}_4 - \vec{r}_3| \cdots |\vec{r}_n - \vec{r}_{n-1}| \cdot |\vec{r}_{n+1} - \vec{r}_n|),$$

$$\nu_{d,n} = \mathbb{E}(|\vec{r}_2 - \vec{r}_1| \cdot |\vec{r}_3 - \vec{r}_2| \cdot |\vec{r}_4 - \vec{r}_3| \cdots |\vec{r}_n - \vec{r}_{n-1}| \cdot |\vec{r}_1 - \vec{r}_n|)$$

? In words, what is the expected product of lengths of an n -step open/closed tour of the points? “Open” means that each point is uniquely visited; “closed” means the same except \vec{r}_{n+1} is replaced by \vec{r}_1 . We do not assume anything about the ordering of the points, hence if $n > d + 1$ the closed tour need not lie on a convex hull boundary.

Mehta & Normand [1] computed

$$\mu_{3,1} = \frac{4}{\sqrt{\pi}} = 2.256758\dots, \quad \mu_{3,2} = 2 + \frac{6\sqrt{3}}{\pi} = 5.307973\dots,$$

$$\mu_{3,3} = \frac{238}{3\sqrt{\pi}} + \frac{56\sqrt{2}}{3\pi^{3/2}} - \frac{216}{\pi^{3/2}} \arctan(\sqrt{2}) = 12.442385\dots,$$

$$\begin{aligned} \mu_{3,4} &= \frac{232}{45} - \frac{3140}{9\pi} + \frac{56}{\sqrt{3}\pi} + \frac{260\sqrt{5}}{9\pi^2} + \frac{912}{\pi^2} \arctan(\sqrt{5}) + \frac{224}{\sqrt{3}\pi^2} \arctan\left(\sqrt{\frac{5}{3}}\right) \\ &= 29.174181\dots, \end{aligned}$$

$$\nu_{3,2} = 6$$

using a multipole expansion (via spherical harmonics) and solution of Fredholm integral equations (via eigenfunction analysis). No exact pattern for either $\mu_{3,n}$ or $\nu_{3,n}$ is observed. It is tempting to believe that the difficulties encountered when $d = 3$

⁰Copyright © 2015 by Steven R. Finch. All rights reserved.

might be somehow circumvented when $d = 1$ or $d = 2$. This turns out to be false. On a line,

$$\begin{aligned}\mu_{1,1} &= \frac{2}{\sqrt{\pi}} = 1.128379\dots, & \mu_{1,2} &= \frac{1}{3} + \frac{2\sqrt{3}}{\pi} = 1.435991\dots, \\ \mu_{1,3} &= \frac{5}{\sqrt{\pi}} + \frac{4\sqrt{2}}{\pi^{3/2}} - \frac{12}{\pi^{3/2}} \arctan\left(\sqrt{2}\right) = 1.778095\dots, \\ \mu_{1,4} &= \frac{2}{15} + \frac{4\sqrt{5}}{\pi^2} + \frac{8}{\pi^2} \arctan\left(\frac{\sqrt{5}}{7}\right) + \frac{8\sqrt{3}}{\pi^2} \arctan\left(\sqrt{\frac{3}{5}}\right) = 2.215483\dots, \\ \nu_{1,2} &= 2, & \nu_{1,3} &= \frac{3}{\sqrt{\pi}} = 1.692568\dots, & \nu_{1,4} &= \frac{2}{3} - \frac{8}{\pi} + \frac{8\sqrt{3}}{\pi} = 2.530818\dots\end{aligned}$$

The formula for $\mu_{1,3}$ in [1] contains typographical errors and that for $\mu_{1,4}$ does not appear at all; we give proofs in Section 2. In the plane,

$$\begin{aligned}\mu_{2,1} &= \sqrt{\pi} = 1.772453\dots, \\ \mu_{2,2} &= 4E\left(\frac{1}{2}\right) - \frac{3}{2}K\left(\frac{1}{2}\right) = 3.341223\dots, \\ \nu_{2,2} &= 4\end{aligned}$$

where

$$\begin{aligned}K(\xi) &= \int_0^{\pi/2} \frac{1}{\sqrt{1 - \xi^2 \sin^2(\theta)}} d\theta = \int_0^1 \frac{1}{\sqrt{(1-t^2)(1-\xi^2 t^2)}} dt, \\ E(\xi) &= \int_0^{\pi/2} \sqrt{1 - \xi^2 \sin^2(\theta)} d\theta = \int_0^1 \sqrt{\frac{1 - \xi^2 t^2}{1 - t^2}} dt\end{aligned}$$

are complete elliptic integrals of the first and second kind [2, 3]. Also [4, 5],

$$\begin{aligned}\nu_{2,3} &= \frac{4}{3\pi} \int_0^\infty \int_0^x \int_{x-y}^{x+y} \frac{x^2 y^2 z^2}{\sqrt{(x+y+z)(-x+y+z)(x-y+z)(x+y-z)}} \\ &\quad \exp\left(-\frac{1}{6}(x^2 + y^2 + z^2)\right) dz dy dx \\ &\approx 6.359,\end{aligned}$$

$$\mu_{2,3} = \lim_{\rho \rightarrow -\frac{1}{2}^+} F(\rho) \approx 6.25$$

and we examine the latter in Section 3. An expression for $F(\rho)$ and explanation of its significance are forthcoming. Finally, returning to three-space [4, 5],

$$\nu_{3,3} = \frac{2\sqrt{3}}{9\pi} \int_0^\infty \int_0^x \int_{x-y}^{x+y} x^2 y^2 z^2 \exp\left(-\frac{1}{6}(x^2 + y^2 + z^2)\right) dz dy dx \approx 12.708.$$

Accurate numerical values for $\mu_{2,4}$, $\nu_{2,4}$, $\nu_{3,4}$ (obtained by procedures other than Monte Carlo simulation) remain at large.

1. ONE-SPACE

Define correlation coefficients

$$\begin{aligned} \rho_{ij} &= \frac{1}{2} \mathbb{E}((r_{i+1} - r_i)(r_{j+1} - r_j)) \\ &= \begin{cases} 1 & \text{if } |i - j| = 0, \\ -1/2 & \text{if } |i - j| = 1, \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

and partial correlation coefficients

$$\begin{aligned} \rho_{ij \cdot k} &= \frac{\rho_{ij} - \rho_{ik}\rho_{jk}}{\sqrt{(1 - \rho_{ik}^2)(1 - \rho_{jk}^2)}}, \\ \rho_{ij \cdot k\ell} &= \frac{\rho_{ij \cdot k} - \rho_{i\ell \cdot k}\rho_{j\ell \cdot k}}{\sqrt{(1 - \rho_{i\ell \cdot k}^2)(1 - \rho_{j\ell \cdot k}^2)}}. \end{aligned}$$

(In statistics [6], $\rho_{ij \cdot k\ell} = 0$ would suggest that $r_{i+1} - r_i$ and $r_{j+1} - r_j$ are conditionally independent, given $r_{k+1} - r_k$ and $r_{\ell+1} - r_\ell$. More precisely, $\rho_{ij \cdot k\ell}$ is the correlation between the residuals for i and j resulting from a linear regression of i with k, ℓ and of j with k, ℓ .) Let R_n be the determinant of $(\rho_{ij})_{1 \leq i, j \leq n}$. It follows from general formulas in [7, 8, 9] that

$$\mu_{1,2} = \frac{4}{\pi} \left(\sqrt{R_2} + \rho_{12} \arcsin(\rho_{12}) \right),$$

$$\begin{aligned} \mu_{1,3} &= \frac{8}{\pi^{3/2}} \left[\sqrt{R_3} + (\rho_{12} + \rho_{13}\rho_{23}) \arcsin(\rho_{12 \cdot 3}) + \right. \\ &\quad \left. (\rho_{13} + \rho_{12}\rho_{23}) \arcsin(\rho_{13 \cdot 2}) + (\rho_{23} + \rho_{12}\rho_{13}) \arcsin(\rho_{23 \cdot 1}) \right], \end{aligned}$$

$$\begin{aligned} \mu_{1,4} = & \frac{16}{\pi^2} \left[\sqrt{R_4} + \sqrt{1 - \rho_{12}^2} (\rho_{34} + \rho_{13}\rho_{14} + \rho_{23}\rho_{24}) \arcsin(\rho_{34,12}) + \right. \\ & \sqrt{1 - \rho_{13}^2} (\rho_{24} + \rho_{12}\rho_{14} + \rho_{23}\rho_{34}) \arcsin(\rho_{24,13}) + \\ & \sqrt{1 - \rho_{14}^2} (\rho_{23} + \rho_{12}\rho_{13} + \rho_{24}\rho_{34}) \arcsin(\rho_{23,14}) + \\ & \sqrt{1 - \rho_{23}^2} (\rho_{14} + \rho_{12}\rho_{24} + \rho_{13}\rho_{34}) \arcsin(\rho_{14,23}) + \\ & \sqrt{1 - \rho_{24}^2} (\rho_{13} + \rho_{12}\rho_{23} + \rho_{14}\rho_{34}) \arcsin(\rho_{13,24}) + \\ & \left. \sqrt{1 - \rho_{34}^2} (\rho_{12} + \rho_{13}\rho_{23} + \rho_{14}\rho_{24}) \arcsin(\rho_{12,34}) \right] + \\ & 4(\rho_{12}\rho_{34} + \rho_{13}\rho_{24} + \rho_{14}\rho_{23}) \gamma \end{aligned}$$

where

$$\gamma = \mathbb{E} [\text{sgn}((r_2 - r_1)(r_3 - r_2)(r_4 - r_3)(r_5 - r_4))]$$

and $\text{sgn}(x) = 1$ when $x \geq 0$; $\text{sgn}(x) = -1$ when $x < 0$. The additional factor γ could be troublesome were it not known that the orthant probability

$$\mathbb{P} \{r_2 - r_1 > 0, r_3 - r_2 > 0, r_4 - r_3 > 0, r_5 - r_4 > 0\} = \frac{1}{16} + \frac{1}{8\pi} \sum_{i < j} \arcsin(\rho_{ij}) + \frac{\gamma}{16}$$

is equal to $1/120$ [10, 11, 12, 13]. From this, we deduce that $\gamma = 2/15$ and thus the value for $\mu_{1,4}$ is confirmed.

2. TWO-SPACE

Thinking of vectors as complex numbers:

$$\vec{r}_2 - \vec{r}_1 = z_1 = (x_2 - x_1) + i(y_2 - y_1), \quad \vec{r}_3 - \vec{r}_2 = z_2 = (x_3 - x_2) + i(y_3 - y_2)$$

we have

$$\bar{z}_1 = (x_2 - x_1) - i(y_2 - y_1), \quad \bar{z}_2 = (x_3 - x_2) - i(y_3 - y_2)$$

and thus

$$\begin{aligned} \psi_{11} &= \mathbb{E}(z_1 \bar{z}_1) = \mathbb{E}(x_2^2 + x_1^2 + y_2^2 + y_1^2) = 4, \\ \psi_{12} &= \mathbb{E}(z_1 \bar{z}_2) = \mathbb{E}(-x_2^2 - y_2^2) = -2. \end{aligned}$$

Writing

$$\Psi = \begin{pmatrix} \psi_{11} & \psi_{12} \\ \bar{\psi}_{12} & \psi_{22} \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix},$$

we obtain inverse covariance matrix

$$\Phi = \begin{pmatrix} \varphi_{11} & \varphi_{12} \\ \bar{\varphi}_{12} & \varphi_{22} \end{pmatrix} = \Psi^{-1} = \begin{pmatrix} 1/3 & 1/6 \\ 1/6 & 1/3 \end{pmatrix}$$

with determinant $\Delta = 1/12$. The joint density of $a = |z_1|$, $b = |z_2|$ is [2, 3]

$$4\Delta a b \exp(-\varphi_{11}a^2 - \varphi_{22}b^2) I_0(2ab|\varphi_{12}|)$$

where I_k is the k^{th} modified Bessel function of the first kind, and therefore

$$\mu_{2,2} = \frac{1}{3} \int_0^\infty \int_0^\infty a^2 b^2 \exp\left(-\frac{1}{3}(a^2 + b^2)\right) I_0\left(\frac{ab}{3}\right) db da = 4E\left(\frac{1}{2}\right) - \frac{3}{2}K\left(\frac{1}{2}\right).$$

The joint density of $a = |z_1|$, $b = |z_2|$, $c = |z_3|$ is more complicated. It is useful to introduce a parameter ρ for which

$$\Psi(\rho) = 4 \begin{pmatrix} 1 & \rho & 0 \\ \rho & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 4 & -2 & 0 \\ -2 & 4 & -2 \\ 0 & -2 & 4 \end{pmatrix}$$

as $\rho \rightarrow -1/2$. The inverse covariance matrix is

$$\Phi(\rho) = \frac{1}{4(1-2\rho^2)} \begin{pmatrix} 1-\rho^2 & -\rho & \rho^2 \\ -\rho & 1 & -\rho \\ \rho^2 & -\rho & 1-\rho^2 \end{pmatrix}$$

and possesses determinant

$$\Delta(\rho) = \frac{1}{64(1-2\rho^2)}.$$

Restrict $-1/\sqrt{2} < \rho < 0$ so that all φ_{ij} and Δ are positive. Let $\varepsilon_0 = 1$, $\varepsilon_k = 2$ for $k \geq 1$. The joint density $f(a, b, c)$ is given by an infinite series [2, 3]

$$\begin{aligned} & 8\Delta a b c \exp(-\varphi_{11}a^2 - \varphi_{22}b^2 - \varphi_{33}c^2) \sum_{k=0}^{\infty} \varepsilon_k (-1)^k. \\ & I_k(2ab|\varphi_{12}|) I_k(2bc|\varphi_{23}|) I_k(2ac|\varphi_{13}|) \\ & = \frac{abc}{8(1-2\rho^2)} \exp\left(-\frac{(1-\rho^2)a^2 + b^2 + (1-\rho^2)c^2}{4(1-2\rho^2)}\right) \sum_{k=0}^{\infty} \varepsilon_k (-1)^k. \\ & I_k\left(\frac{-\rho ab}{2(1-2\rho^2)}\right) I_k\left(\frac{-\rho bc}{2(1-2\rho^2)}\right) I_k\left(\frac{\rho^2 ac}{2(1-2\rho^2)}\right) \end{aligned}$$

which is evidently divergent when $\rho = -1/2$. Convergence seems to stabilize for ρ slightly to the right of $-1/2$. We therefore examine

$$\mu_{2,3} = \lim_{\rho \rightarrow -\frac{1}{2}^+} \int_0^\infty \int_0^\infty \int_0^\infty a b c f(a, b, c) dc db da \approx 6.25,$$

yielding a result consistent with (but unfortunately not improving upon) computer simulation.

A higher-dimensional analog of $f(a, b, c)$ is exhibited in [14, 15] but requires that the 4×4 inverse covariance matrix have corner entry $\varphi_{14} = 0$. This is not true in our scenario. More relevant discussion is found in [16, 17, 18, 19].

REFERENCES

- [1] M. L. Mehta and J.-M. Normand, On some definite multiple integrals, *J. Phys. A* 30 (1997) 8671–8684; MR1619533 (99g:33059).
- [2] K. S. Miller, Complex Gaussian processes, *SIAM Rev.* 11 (1969) 544–567; MR0258109 (41 #2756).
- [3] K. S. Miller, *Complex Stochastic Processes. An Introduction to Theory and Application*, Addison-Wesley, 1974, pp. 86–100; MR0368118 (51 #4360).
- [4] P. Clifford and N. J. B. Green, Distances in Gaussian point sets, *Math. Proc. Cambridge Philos. Soc.* 97 (1985) 515–524; MR0778687 (86i:62091).
- [5] S. R. Finch, Random triangles, unpublished note (2010), <http://www.people.fas.harvard.edu/~sfinch/>.
- [6] K. Baba, R. Shibata and M. Sibuya, Partial correlation and conditional correlation as measures of conditional independence, *Austral. New Zealand J. Statist.* 46 (2004) 657–664; MR2115961 (2005k:62153).
- [7] S. Nabeya, Absolute moments in 2-dimensional normal distribution, *Annals Inst. Statist. Math.* 3 (1951) 2–6; MR0045347 (13,570b).
- [8] S. Nabeya, Absolute moments in 3-dimensional normal distribution, *Annals Inst. Statist. Math.* 4 (1952) 15–30; MR0052072 (14,569c).
- [9] S. Nabeya, Absolute and incomplete moments of the multivariate normal distribution, *Biometrika* 48 (1961) 77–84; MR0126917 (23 #A4211).
- [10] R. L. Plackett, A reduction formula for normal multivariate integrals, *Biometrika* 41 (1954) 351–360; MR0065047 (16,377c).

- [11] J. A. McFadden, Two expansions for the quadrivariate normal integral, *Biometrika* 47 (1960) 325–333; MR0119221 (22 #9987).
- [12] D. R. Childs, Reduction of the multivariate normal integral to characteristic form, *Biometrika* 54 (1967) 293–300; MR0214177 (35 #5028).
- [13] H.-J. Sun and C. Asano, On the normal orthant probability with a tri-diagonal correlation matrix, *Eng. Sci. Rep. Kyushu Univ.* 12 (1990) 53–58.
- [14] L. E. Blumenson and K. S. Miller, Properties of generalized Rayleigh distributions, *Annals Math. Statist.* 34 (1963) 903–910; MR0150860 (27 #846).
- [15] Y. Chen and C. Tellambura, Infinite series representations of the trivariate and quadrivariate Rayleigh distribution and their applications, *IEEE Trans. Comm.* 53 (2005) 2092–2101.
- [16] S. Nadarajah and S. Kotz, On the infinite series representations for multivariate Rayleigh distributions, *IEEE Trans. Comm.* 55 (2007) 392–393.
- [17] S. Nadarajah and S. Kotz, Comments on “A trivariate chi-squared distribution derived from the complex Wishart distribution”, *J. Multivariate Anal.* 99 (2008) 306–307; MR2432328.
- [18] W. V. Li and A. Wei, Gaussian integrals involving absolute value functions, *High Dimensional Probability V*, Proc. 2008 Luminy conf., ed. C. Houdré, V. Koltchinskii, D. M. Mason and M. Peligrad, Inst. Math. Statist., 2009, pp. 43–59; MR2797939 (2012f:60059).
- [19] W. V. Li and A. Wei, A Gaussian inequality for expected absolute products, *J. Theoret. Probab.* 25 (2012) 92–99; MR2886380.

Steven Finch
 Dept. of Statistics
 Harvard University
 Cambridge, MA, USA
steven_finch@harvard.edu